

Are spatial inequalities growing? The scale of population concentrations in England and Wales

Environment and Planning A
2016, Vol. 48(7) 1318–1336



© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0308518X15621306

epn.sagepub.com



Christopher D Lloyd

Department of Geography and Planning, School of Environmental Sciences, University of Liverpool,
Liverpool, UK

Abstract

This paper explores how the population of England and Wales in 2001 and in 2011 was spatially concentrated by a range of demographic, social and economic characteristics. Where members of population sub-groups tend to live apart from members of other sub-groups then the population may be regarded as geographically unequal. In the UK, debates about the north–south divide have reflected the principal geographical division in public perception, with wealth and health inequalities at the forefront. This analysis uses variograms to characterise the differences between areas over multiple spatial scales. There is evidence for stronger spatial structure (more distinct spatial patterning) in variables including car and van availability and ethnicity than in age, self-reported illness, and qualifications, and these relate to urban–rural differences in the former variables. The key contribution of the paper is in using directional variograms and variogram maps to show marked differences in population concentrations by direction with, for example, north–south differences in qualifications being (on average) greater than those in the east–west direction. However, for most variables which show increased variation (and thus suggest increased geographical inequalities) between 2001 and 2011, increases are proportionally similar in all directions. Only in the case of self-reported ill-health does the north–south (or, in this case, north west–south east) divide appear to have increased.

Keywords

Inequalities, segregation, census, variogram

Introduction

This paper explores how the population of England and Wales in 2001 and in 2011 was spatially concentrated by a range of demographic, social and economic characteristics. Where members of a population sub-group tend to live apart from members of another

Corresponding author:

Christopher D Lloyd, Department of Geography and Planning, School of Environmental Sciences, University of Liverpool,
Roxby Building, Chatham Street, Liverpool L69 7ZT, UK.
Email: c.d.lloyd@liverpool.ac.uk

sub-group then the population may be regarded as geographically unequal in some sense. The paper shows how variograms can be used to characterise geographic inequalities between regions (for example, are neighbourhoods with low rates of qualifications becoming more geographically concentrated?). This analysis is extended using directional variograms and variogram maps which can be used to assess, for example, if the north and south of a study area tend to be more different to one another than the east and west. Directional variograms allow for the assessment of the spatial scale of population concentrations in different directions. In other words, it is possible to determine if geographic inequalities are increasing at a relatively small spatial scale (e.g. a city scale) or a larger scale (e.g. between regions within a country) and the directions in which these inequalities are changing most (e.g. the north and south may be growing apart, while the east and west are becoming more alike). The paper provides an analytical framework for exploring the complexities of geographic inequalities in any context where small area population data are available.

In the UK, debates about the north–south divide have reflected the principal division in public perception, with differences in wealth and health emphasised. For example, Doran et al. (2004) identified a north west–south east divide in social class inequalities in Britain in 2001 with higher rates of poor health in Wales, the North East and North West regions of England than elsewhere, but with the widest health gaps between social classes in Scotland and London. However, Bland (2004), in a response to the article, argued that the differences between socio-economic groups do not vary markedly between regions. Dorling and Thomas (2004), in a review of geographical patterns in the population of the UK in 2001, state that there are increasing regional inequalities between London and the South East and the rest of the country in terms of educational qualifications. These growing inequalities are a function, they argue, of differential flows of people with different levels of educational qualifications. Many other studies have identified north–south divides and Hacking et al. (2011) observe substantial inequalities in all-cause mortality between the North and South of England between 1965 and 2008, and they note an increase in inequalities between 2000 and 2008. Gardiner et al. (2013) are focused instead on spatially unbalanced growth in the British economy and they argue that the north–south economic divide has become more marked in recent decades.

In other studies which consider geographical inequalities, the north–south divide may not be a direct focus. Thomas et al. (2010) argue that inequalities in age–sex standardised mortality ratios (aged under 75) in Britain have increased every two years from 1990–1991 to 2006–2007. In a study focusing on earnings inequalities, Stewart (2011) argues that there was a growth in overall inequality between 1997 and 2008 and that this was mostly driven by London and by the financial sector, with smaller increases in the South East and East Anglia, while changes in the rest of Britain were very small. Fahmy et al. (2011) find that poverty at the household level increased markedly between 1971 and 2011 and it also became more spatially concentrated during the period. Another focus has been on differences between urban and rural areas; Riva et al. (2011) assessed the role of residential mobility in explaining geographic inequalities in all-cause mortality between urban and rural areas. They found that residential mobility between 1981 and 2001 explained about 30% of the urban–rural inequalities in mortality observed at the end of this period. Tunstall (2011) demonstrates that, for those in social housing in England, some dimensions of social exclusion (income, employment, and neighbourhood quality) had reduced by a small amount between 2000 and 2011, but there was an indication of an increase in concentrations of disability.

Most previous studies which seek to assess evidence for geographical divides in the UK tend to be limited by being based on data for regions, and differences are usually considered

in terms of predefined sets of regions which collectively comprise, for example, 'the North' and 'the South'. An arguably better approach is to instead assess differences in a range of directions and thus to use the data in a flexible way to determine how far differences between areas may be greater in some directions than in others. A second limitation of many studies is that they are generally based on data for large areas, whereas a geographically refined approach might be more revealing. This paper offers a solution to both of these limitations: it is based on data for the smallest Census output geographies (output areas; OAs) and it uses directional variograms as a means of characterising differences between small areas (which equate, for some variables at least, to geographical inequalities) in several directions. Using these data and this approach, it is possible to explore the *spatial scale* and *magnitude* of differences between small areas by direction without making a prior assumption about which directions (e.g. north–south, east–west...) may represent the main contrasts in terms of their population characteristics. Use of zones larger than OAs would impact on the ability to measure the magnitude of differences and so a strong case can be made for using the smallest available zones. This is the first piece of published research which has sought to examine inter-regional differences in such a way and it provides an original approach to the analysis of geographical inequalities and enhances our understanding of the demographic, social and economic geographies of England and Wales and the ways in which they changed between 2001 and 2011.

The present analysis makes use of counts for OAs by age, ethnic group, housing tenure, car or van availability, qualifications, employment, limiting long-term illness (LLTI) and National Statistics Socio-economic Classification (NS-SEC). Geographical inequalities link directly to research on residential segregation where the objective is to assess how members of different population groups may live together or apart (see Lloyd et al., 2014 for a recent overview). Massey and Denton (1988) assess alternative dimensions of segregation, while more recent research has focused on explicitly spatial measures of segregation (e.g. Brown and Chung, 2006; Reardon et al., 2008, 2009; Wong, 2004). Many studies of segregation have focused on urban areas and they have had a particular concern with ethnic segregation (e.g. Johnston et al., 2007; Catney, 2015, who focuses on segregation in both urban areas and in England and Wales as a whole) and socio-economic segregation (e.g. Quillian and Lagrange, 2013; Musterd, 2005, considers both ethnic and socio-economic segregation within European cities). The particular focus in the present study is on characterising how population sub-groups are spatially concentrated – for example, how large, on average, are differences within urban areas as opposed to differences between areas? Age categories are included in the analysis as, although age may be less strongly linked to inequality than, for example, (un)employment, spatial patterns in the population by age are important as an indicator of population change.

This paper builds on research conducted by, among others, Voas and Williamson (2000), Dorling and Rees (2003) and Lloyd (2015). Voas and Williamson (2000) sought to assess unevenness (measured using the index of dissimilarity, *D*) in population groups using 1991 Census data for England and Wales. Using data for districts, wards and enumeration districts, they considered what proportion of unevenness was found at each of the three scales represented by these zones. Dorling and Rees (2003) explored changes in spatial divisions across Britain between 1971 and 2001. In the study by Lloyd (2015), the spatial structure of population sub-groups in England and Wales in 2001 and 2011 was the focus and the analyses suggested that differences between regions, in terms of most population sub-groups considered, had reduced between 2001 and 2011. Other studies have focused on changes by ethnicity (e.g. Catney, 2013) and demographic and deprivation change

(Norman, 2010). The present study assesses systematically, for the first time, the scales over which population sub-groups are distributed across England and Wales. Resource allocation by government depends on knowledge of the scales over which population sub-groups (for example, those with poor health) are distributed and so the analyses set out in this paper could have relevance for policy. However, while area-based interventions are commonly proposed (e.g. Smith, 1999), it is worth noting that some authors question their value as a primary means of tackling poverty. Townsend (1979) argues that any socially or economically deprived areas that are identified are likely to exclude many (and maybe most) of those in poverty while these 'deprived areas' will include many people who are not deprived, a theme which links to the well-known ecological fallacy. Despite such criticisms, it is argued here that greater knowledge of how population groups are geographically unequal is vital to better understanding inequalities and in combating them.

The paper provides a systematic analysis of spatial scale of variation in a set of population sub-groups by computing variograms for each variable for 2001 and 2011, building on the examples given by Lloyd (2015). In addition, it considers differences in spatial variation in population sub-groups by direction. Firstly, the data used in the analysis are described and the computation of log-ratios (derived from percentages of people in particular sub-groups), which provide the inputs for the analyses, is summarised. Next, the analysis of spatial variation with the variogram is summarised. The analyses provide a rich picture of spatial variation in the selected population sub-groups in England and Wales in 2001 and 2011.

Data processing

The data

The analysis is based on counts reported for OAs, the smallest areas for which counts from the UK Census are available. The focus is on England and Wales rather than the UK as a whole since the analysis builds on previous research (Lloyd, 2015). There were 175,434 OAs in England and Wales in 2001 (mean population = 297), while in 2011 there were 181,408 (mean population = 309). Only some 2.6 per cent of 2001 OAs were changed as a result of the 2011 Census¹ and results based on the two sets of OAs were considered to be comparable. The small cell adjustment procedure for 2001 counts was judged to have a minimal impact for national-level analyses of generally large population groups (see Williamson, 2007, for discussion about the impact of these procedures). OAs were first used in the 2001 Census in England and Wales and they were constructed using clusters of adjacent unit postcodes; they were intended to have similar population sizes and to be as socially homogenous as possible according to tenure of household and dwelling type.² The automated zone design methodology used to generate OAs is detailed by Martin et al. (2001).

The data used in the current analysis, as specified in Table 1, are from the Key Statistics tables. The aim was to select an array of demographic, social and economic variables which are important components of population profiles in small areas. For example, age profiles, (un)employment and health status are key population characteristics. The range of variables and specific groupings (e.g. sets of age ranges) are necessarily limited to make this exploratory analysis manageable. In addition, the variables represent very different patterns of spatial variability and thus they enable a proper assessment of the selected methods. Figure 1 includes maps of two selected variables for context: the percentages of

Table 1. Key statistics census tables and derived variables and counts.

2001	2011	Table	Variable	Abbreviation	Definition	2001 count	2011 count
KS001	KS102	Age structure	Age 0 to 15	A0to15	Persons aged 0 to 15	10,488,736	10,579,132
			Age 16 to 29	A16to29	Persons aged 16 to 29	9,112,810	10,495,245
			Age 30 to 64	A30to64	Persons aged 30 to 64	24,127,596	25,778,462
			Age 65 plus	A65plus	Persons aged 65 plus	8,312,774	9,223,073
KS006	KS201	Ethnic group	All Whites	White	White persons	47,520,866	48,209,395
			Non-Whites	Non-White	Non-White persons	4,521,050	7,866,517
KS018	KS402	Housing tenure	Owner occupied	OwnOcc	Owner occupied HH	14,916,465	15,031,914
			Social rented	SocRent	Social rented HH	4,157,251	4,118,461
			Private rented	PrivRent	Private rented HH	2,586,759	4,215,669
KS017	KS404	Cars and vans	Cars or vans	CarsVans	HH with cars or vans	15,858,292	17,376,274
			No cars or vans	NoCarsVans	HH with no cars or vans	5,802,183	5,989,770
KS013	KS501	Qualifications and students	Qualifications	Qual	Persons with qualifications	26,670,396	35,189,453
			No qualifications	NoQual	Persons with no qualifications	10,937,042	10,307,327
KS09A	KS601	Economic activity – all persons (aged 16–74)	Employed economically active	EAEmploy	EA employed persons	22,795,520	25,449,863
			Unemployed	EAUnemp	EA unemployed persons	1,261,343	1,799,536
KS14A	KS611	National Statistics Socio-economic Classification (NS-SEC; aged 16–74) ^a	NS-SEC 1, 2	NS-SEC12	NS-SEC12	10,172,697	12,792,224
			NS-SEC 3 to 7	NS-SEC37	NS-SEC3–7	16,650,975	22,324,839
			NS-SEC 8	NS-SEC8	NS-SEC8	1,404,188	2,301,614
KS008	KS301	Health and provision of unpaid care	No LLTI	No LLTI	Persons with no LLTI	42,557,060	46,027,471
			LLTI	LLTI	Persons with a LLTI	9,484,856	10,048,441

Note: Counts for 2001 are England and Wales level counts and they differ from the sums of the OA-level counts because of small cell adjustment.

PrivRent for 2011 includes 'Private rented: Private landlord or letting agency', 'Private rented: Other' and 'Living rent free'. Qual and NoQual figures for 2001 and 2011 (in italics) use 16–74 and 16 plus population bases, respectively, and so should not be directly compared; NS-SEC counts for 2011 (in italics) include imputed persons.

HH: households; EA: economically active; LLTI: limiting long-term illness.

^aThe NS-SEC classes are as follows: NS-SEC 1.2: Managerial, administrative and professional occupations; NS-SEC 3–7: Intermediate, routine and manual occupations; NS-SEC 8: Never worked and long-term unemployed.

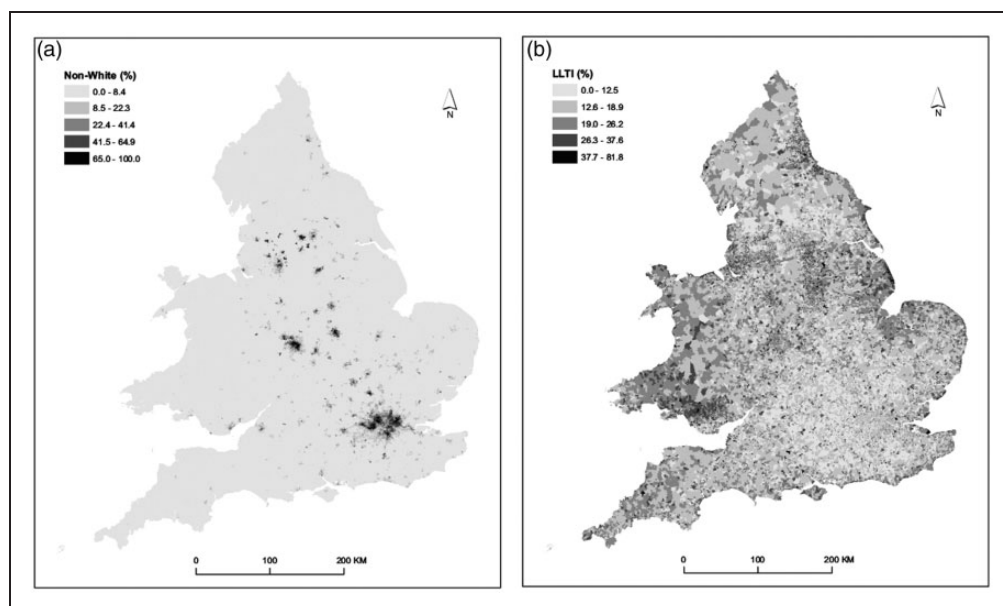


Figure 1. Percentage of (a) non-White persons and (b) persons with a limiting long-term illness (LLTI) by OAs for 2011. Contains national statistics data © Crown copyright and database right 2012. Contains ordnance survey data © Crown copyright and database right 2012.

(a) non-White persons and (b) persons with a LLTI by OAs in 2011. A distinct urban–rural pattern is apparent in the map of non-white persons (Figure 1(a)). Examination of the LLTI map suggests that there are generally lower levels of LLTI in the South than in the North and West. The analyses presented below require information on the distances between OAs and these were measured using the population weighted centroids (median centre of OAs based on household locations and populations; units are British National Grid coordinates in metres) provided by the Office for National Statistics (ONS) (note that the Eastings and Northings of the population weighted centroids were provided as integers). The centroids were considered to be suitable approximations of OA locations given the small sizes of these zones.

Deriving log-ratios

The analysis is based on variograms computed from log-ratio transformed percentages. Percentages are constrained to sum to 100 (while proportions sum to one), and such data are referred to as compositional (Aitchison, 1986). Many studies have argued that statistical analysis of raw percentages or proportions is not appropriate, and transforming percentages or proportions to log-ratios provides one solution (see Lloyd et al., 2012 for an introduction to the topic in a population studies context while Filzmoser et al., 2009, discuss analysis of univariate compositional data). The present analysis is based on balances, a form of isometric log-ratios (ilr) (Egozcue et al., 2003; Egozcue and Pawlowsky-Glahn 2005, 2006). The compositions (sets of percentages) used in the study are two-part (ethnicity, cars and vans, qualifications, employment and LLTI all comprise two sets of percentages), three-part (housing tenure and NS-SEC comprise three sets of percentages)

and four-part (age); the log-ratios were computed as follows (see Table 1 for definitions of input variables):

Two-part compositions

$$\begin{aligned}\text{Ethnicity} &= \sqrt{\frac{1}{2}} \ln \frac{\text{White}}{\text{Non-White}}, \\ \text{CarsVans} &= \sqrt{\frac{1}{2}} \ln \frac{\text{NoCarsVans}}{\text{CarsVans}}, \\ \text{Qual} &= \sqrt{\frac{1}{2}} \ln \frac{\text{NoQual}}{\text{Qual}}, \\ \text{Employ} &= \sqrt{\frac{1}{2}} \ln \frac{\text{EAEmploy}}{\text{EAUnemp}}, \\ \text{LLTI} &= \sqrt{\frac{1}{2}} \ln \frac{\text{LLTI}}{\text{NoLLTI}}\end{aligned}$$

Three-part compositions

$$\begin{aligned}\text{Tenure1} &= \sqrt{\frac{2}{3}} \ln \frac{(\text{OwnOcc} \times \text{PrivRent})^{\frac{1}{2}}}{\text{SocRent}}, \\ \text{Tenure2} &= \sqrt{\frac{1}{2}} \ln \frac{\text{OwnOcc}}{\text{PrivRent}}, \\ \text{NSSEC1} &= \sqrt{\frac{2}{3}} \ln \frac{(\text{NSSEC12} \times \text{NSSEC37})^{\frac{1}{2}}}{\text{NSSEC8}}, \\ \text{NSSEC2} &= \sqrt{\frac{1}{2}} \ln \frac{\text{NSSEC12}}{\text{NSSEC37}}\end{aligned}$$

Four-part compositions

$$\begin{aligned}\text{Age1} &= \sqrt{\frac{3}{4}} \ln \frac{(\text{A0to15} \times \text{A16to29} \times \text{A30to64})^{\frac{1}{3}}}{\text{A65plus}}, \\ \text{Age2} &= \sqrt{\frac{2}{3}} \ln \frac{(\text{A0to15} \times \text{A16to29})^{\frac{1}{2}}}{\text{A30to64}}, \\ \text{Age3} &= \sqrt{\frac{1}{2}} \ln \frac{\text{A0to15}}{\text{A16to29}}\end{aligned}$$

Some counts are zeroes and the percentages were calculated from counts x_1, x_2, x_3, \dots with $x_1 + 1, x_2 + 1, x_3 + 1 \dots$ (Lloyd, 2010 provides justification for such an approach). So, a value of one is added to all counts and the percentages, $y_1, y_2, y_3 \dots$ are calculated from the modified counts. Lloyd (2015) assessed the sensitivity of results to the addition of different values (e.g. 0.1 and 0.5) and the results were found to be robust.

Characterising spatial variation with the variogram

The paper uses the variogram to consider how the population sub-groups are spatially structured over multiple scales. The variogram, $\gamma(\mathbf{h})$, relates half the average of the squared differences (the semivariances) between zones to the distances (in lags or bins)

separating their centroids (here, population weighted centroids). The variogram provides a summary of spatial dependence at different spatial scales. The experimental variogram can be estimated for the $p(\mathbf{h})$ paired observations (log-ratios in the present study), $z(\mathbf{s}_i)$, $z(\mathbf{s}_i + \mathbf{h})$, $i = 1, 2, \dots, p(\mathbf{h})$ with

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2p(\mathbf{h})} \sum_{i=1}^{p(\mathbf{h})} \{z(\mathbf{s}_i) - z(\mathbf{s}_i + \mathbf{h})\}^2 \quad (1)$$

where \mathbf{h} is the lag (distance and direction) by which two observations are separated; so, the paired semivariances are binned (e.g. the average semivariance for all pairs separated by between 2 km and 4 km). Variograms can be estimated using paired data which are aligned in a particular direction. In the case of irregularly distributed data (e.g. OA centroids), pairs can be compared based on a predefined lag distance, direction and angular tolerance and this is discussed further under ‘Directional variograms and variogram maps’.

The form of the variogram is obviously a function of the data support (the geometrical size, shape and orientation of the zones for which data values are reported; see Lloyd, 2014), and, in the present analysis, variograms are computed from log-ratios derived from data for OAs for 2001 and 2011. Goovaerts (2008) outlines an approach for deriving the point support variogram from area data. The analysis of compositional data using variograms is discussed by Pawlowsky and Burger (1992) and Pawlowsky-Glahn and Olea (2004). The weighting of variograms by population has been assessed by some researchers (see Goovaerts et al., 2005, and was tested by Lloyd, 2015); only unweighted variograms are used here since OAs are constrained to have a minimum population size and increasing the influence of those with larger populations was not considered desirable in this case.

Models are often fitted to variograms (particularly as a input to kriging interpolation; Webster and Oliver, 2007). Fitted models also provide a useful summary of the structure of variograms and models are fitted to the variograms estimated from OA data. The models fitted in this analysis comprise two elements – a nugget effect and one or two spherical model components. A spherical model component is defined by the range (denoted by a , representing the spatial scale of variation) and the structured component (c , representing spatially correlated variation); more than one spherical model component is fitted to some variograms which have more complex forms. The nugget effect, c_0 , represents measurement error and variation at a distance smaller than that represented by the sample spacing. The nugget effect plus the structured component(s) is the total sill (the *a priori* variance). The range indicates the spatial *scale* of variation while the nugget effect and structured component(s) indicate the *magnitude* of variation. An example model is described below with reference to the nugget effect, ranges and structured components. The variograms were estimated using the Gstat software (Pebesma and Wesseling, 1998). The models were fitted in Gstat using weighted least squares, whereby the weights are a function of the number of paired observations at each lag. The variograms are here estimated from over 175,000 observations for both years and all semivariances (for both omnidirectional and directional variograms) are estimated from tens of millions of pairs. Thus, assessments of changes in variograms between years are considered robust (see Webster and Oliver, 2007, for discussion about variograms and confidence intervals derived as part of an analysis based on simulated data).

Context: Spatial patterning in population sub-groups

Prior to detailing the analysis of log-ratios using variograms, it is useful to consider some expected patterns given what is known about the spatial distribution of population

sub-groups in England and Wales. Dorling and Thomas (2004) provide maps and discussions around key population variables in 2001 along with analyses of change since 1991 and these, along with maps generated for the current analysis, provide a starting point for interpretation of the variograms. Taking each of the variables considered in turn – Age: there are more young adults in London and other major cities and towns than elsewhere; Ethnicity: immigrant settlement areas shape the ethnic geographies of England and Wales (see Catney and Simpson 2010); Tenure: social rented households tend to be more proportionately numerous in urban areas, especially London; there are lower rates of access to cars and vans in urban areas (London and northern cities) and in the Welsh valleys than in many other areas; Dorling and Thomas (2004) note a marked north–south divide in qualifications with higher rates of no qualifications in the North and West than in the South and East; Employment: there tend to be higher rates of unemployment in the North and West than in South and East (with the exception of London); NS-SEC: there is evidence for a pronounced geography of occupational classifications with, for example, high rates of professionals and managers in London and other parts of the South East; Poor health: Dorling and Thomas found that rates of poor health were low in the South East with the exception of London with high rates in the Welsh valleys and urban areas of Northern England. These observations are revisited in the following sections detailing the analysis of spatial variations in population sub-groups using variograms. Differences in the nature of spatial patterns (e.g., broad regional trends or urban–rural contrasts) will relate to different variogram forms and this is elucidated below.

Omnidirectional variograms

The focus is first on spatial variation in all directions – no account is taken of, for example, east–west or north–south differences. In this case, relatively small distances are considered. The second part of the analysis explores variation in different directions and over larger distances with the aim of determining how far there are regional trends in population sub-group concentrations and assessing how far these trends may have changed between 2001 and 2011.

Variograms for each set of log-ratios (2 km lag to a maximum of 80 km) for OAs for 2001 and 2011 were computed and six of these were selected for display (a subset being necessary for reasons of space); they are shown in Figure 2 with models fitted in each case (in some cases the model does not extend to the maximum lag – this indicates it was only fitted to some subset of the semivariances). The variograms selected for display were for Ethnicity, Qual, Tenure1, Tenure2, Employ and LLTI and these were picked because of the contrasting directional variation they displayed, as discussed in the following section. Table 2 summarises the nugget effects, structured components and the ranges of the models fitted to all 12 variograms (not just for the selected six). The widely applied spherical model (see Webster and Oliver, 2007) was used in all cases for ease of comparison. As an example, the variogram for Ethnicity in 2011 has a nugget effect (c_0) of 0.172; for the first spherical component, it has a structured component (c_1) of 0.35 and a range (a_1) of 6263 m while, for the second spherical component, it has a structured component (c_2) of 0.69 and a range (a_2) of 56304 m. The range values indicate dominant spatial features at a local scale (the first range; approximately 6 km) and over a larger ‘regional’ scale (with a range of approximately 56 km). Smaller range figures would indicate more localised patterns, while larger range figures suggest more ‘gradual’ spatial trends. Note that distances are measured directly from the coordinates and so they do not account for the curvature of the Earth. However, differences in distances were considered of negligible importance in this context given that approximate distances would be sufficient to identify major scales of variation.

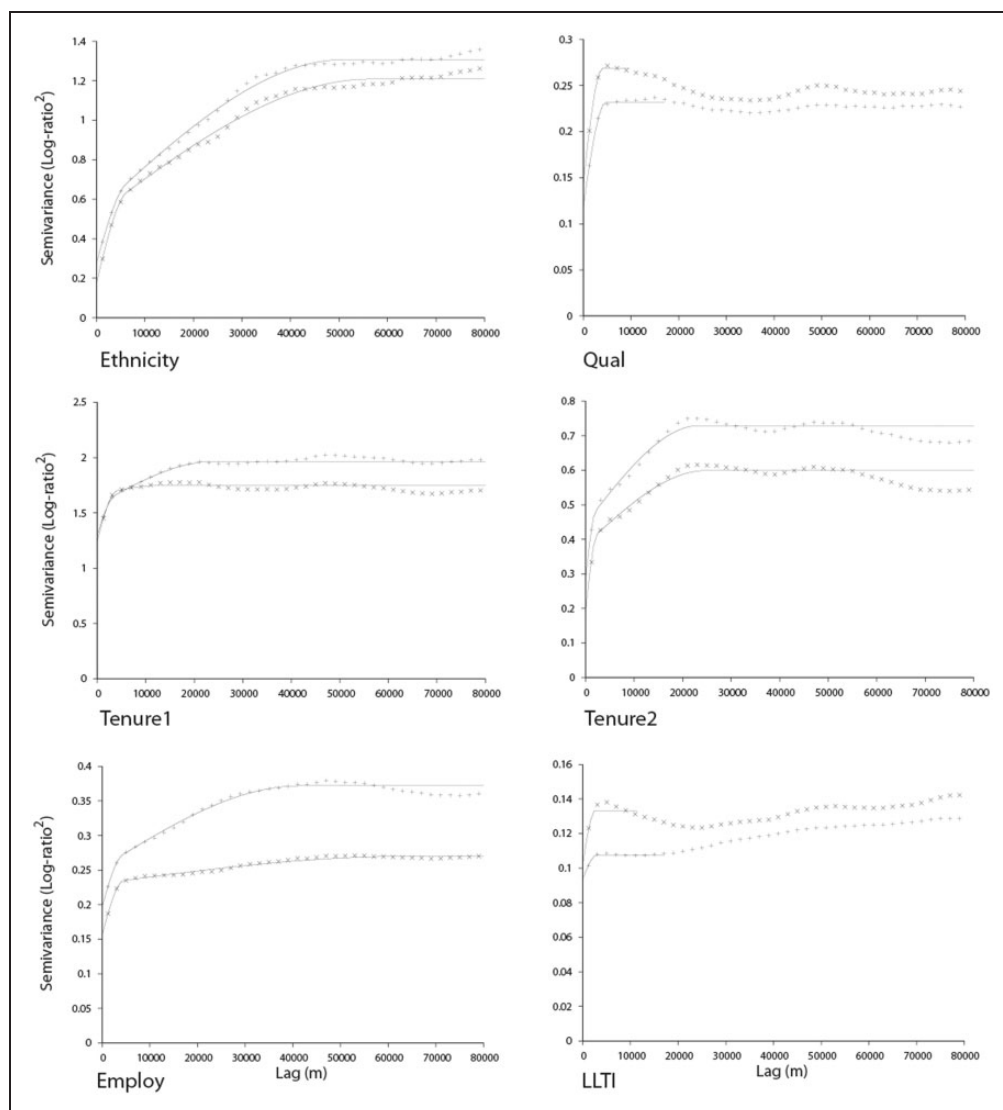


Figure 2. Variograms for selected log-ratios for OAs (2 km lag).

Thus, log-ratios which have short-range spatial variation (e.g. Age2, Age3 (the variograms for age are not shown), Qual and LLTI, for which there is little spatial structure over larger areas), can be distinguished from those which have much longer range spatial variation (e.g. Ethnicity and CarsVans (not shown), highlighting the contrast between urban and rural areas) and the range could be seen as a measure of the scale over which a log-ratio is spatially concentrated. There is evidence for very large-scale trends in Qual and LLTI (see ‘Directional variograms and variogram maps’), but little indication of distinct regional-scale concentrations. The forms of the variograms correspond broadly to what might be expected given the observations under ‘Directional variograms and variogram maps’ (e.g. strong evidence of spatial structure in the cases of variables which exhibit clear urban–rural contrasts) but quantitative summaries of the spatial scale and magnitude of variation, as provided by variograms, constitute an objective means of comparatively

Table 2. Moran's I and standard deviations of log-ratios; variogram model coefficients.

Year	Variable	SD	c_0	c_1	Model	a_1	c_2	Model	a_2
2001	Age1	0.692	0.352	0.033	sph	3679.43			
2011	Age1	0.731	0.344	0.087	sph	4263.96	0.113	sph	86839.5
2001	Age2	0.301	0.058	0.024	sph	3243.12			
2011	Age2	0.294	0.048	0.03	sph	2621.18			
2001	Age3	0.411	0.098	0.102	sph	2560			
2011	Age3	0.418	0.089	0.129	sph	2773.67			
2001	Ethnicity	1.168	0.275	0.266	sph	6101.16	0.766	sph	51013.3
2011	Ethnicity	1.133	0.172	0.35	sph	6262.94	0.686	sph	56303.7
2001	Tenure1	1.409	1.304	0.271	sph	3576.57	0.395	sph	24080.3
2011	Tenure1	1.309	1.251	0.412	sph	4014.59	0.087	sph	12878.5
2001	Tenure2	0.832	0.282	0.163	sph	2073.48	0.284	sph	23406.4
2011	Tenure2	0.739	0.189	0.196	sph	2369.04	0.214	sph	24782.4
2001	CarsVans	0.776	0.264	0.378	sph	30782			
2011	CarsVans	0.776	0.243	0.421	sph	26846.5			
2001	Qual	0.493	0.119	0.113	sph	4740.47			
2011	Qual	0.507	0.145	0.124	sph	4011.36			
2001	Employ	0.602	0.196	0.058	sph	4247.5	0.118	sph	41530.9
2011	Employ	0.522	0.155	0.076	sph	4538	0.039	sph	63928.9
2001	NS-SEC1	0.939	0.444	0.157	sph	4438.73	0.325	sph	40446.5
2011	NS-SEC1	0.771	0.265	0.194	sph	4645.67	0.159	sph	48765.6
2001	NS-SEC2	0.499	0.097	0.088	sph	4424.6	0.035	sph	13018.8
2011	NS-SEC2	0.479	0.089	0.096	sph	4242.29	0.049	sph	10735.9
2001	LLTI	0.371	0.094	0.014	sph	3009.74			
2011	LLTI	0.388	0.102	0.031	sph	2621.18			

SD: standard deviation; sph: spherical.

assessing the geography of different variables and different time points in a way which is almost impossible via visual assessment of maps.

The largest proportional changes in sill values between 2001 and 2011 were for Employ and NS-SEC1 (not shown) (the two sets of log-ratios having similar spatial structures, which is no surprise given that NS-SEC1 relates those with an occupational classification to those who have never worked or are long-term employed). In both of these cases, the sills (and thus the amount of variation) have decreased. Decreases in geographic inequalities can be considered in light of the global recession of 2008–2009 which might have been expected to increase geographical inequalities, rather than reduce them. One possible interpretation is that during, or after, a recession it could appear that inequality has decreased because unemployment has risen in areas where previously it had been very low. Reduction in the Tenure2 variogram sill between 2001 and 2011 reflects less spatial variation in the ratio of owner occupied households to private rented households in 2011 than in 2001. This is likely to be, in part, a function of high house prices, low wage growth and tighter lending requirements (ONS, 2013) which appear to have acted to differentially increase private renting rates making owner occupation rates more similar, on average, between areas. Reduction in the sills for the Ethnicity variograms probably reflect in part, the destinations of migrants from the 2004 European Union accession countries (Lloyd, 2015), as well as dispersal from immigrant settlement areas by members of ethnic minority groups (Catney and Simpson, 2010; Simpson and Finney, 2009).

Directional variograms and variogram maps

The focus now moves onto how far the spatial structure of the log-ratios differs with direction. In variogram estimation (see equation (1)), observations $z(s_i)$ and $z(s_j)$ may be included in calculations if location s_j is north of location (indicated by 0° in clockwise from north) s_i within an angular tolerance of 22.5 decimal degrees. Using an angular tolerance of 22.5 decimal degrees (22.5° either side of the directional lines), one strategy would be to compute directional variograms for 0° , 45° , 90° and 135° , thus giving complete coverage and with no overlap between the directions (Webster and Oliver, 2007). The selection of paired data is illustrated graphically in Figure 3(a) where the specified direction is 45° clockwise from north and the angular tolerance is 22.5° (that is, 22.5° either side of the 45° directional line), in this case $z(s_j)$ would be paired with $z(s_i)$ since it is within the specified tolerance. While omnidirectional variograms characterise the spatial scale of variation and the amount of variation in all directions simultaneously, directional variograms allow for a nuanced assessment of how population characteristics vary by direction. As an example, it is possible to determine if differences in levels of self-reported ill-health between areas of the East and West of England and Wales tend to be smaller or larger than equivalent differences between areas in the North and South. A key contribution of directional variograms is that the scale of these differences can be ascertained – the effect of distances on differences between the North and South may be greater than the effect of distances on differences between the East and West, with paired areas in the East and West tending to be similar even over large distances while only relatively small distances may correspond to large differences between paired areas in the North and South. As an illustrative case of population change, it could be that differences in the share of a population sub-group appear to be constant but, in fact, they have been decreasing in an east–west direction (small areas in the East and West have become, on average, more similar to one another) while they have been increasing in an north–south direction (small areas in the North and South have become, on average, more dissimilar to one another). Directional variograms allow the assessment of changes which vary by direction and a rare example of their application in population studies is the paper by Balabdaoui et al. (2001), who explore changes in spatial patterns in a fertility index in India.

Figure 3(b) provides as an example directional variograms for 2001 and 2011 for the LLTI log-ratios given a 10 km lag, to a maximum of 300 km; the variograms are estimated within tolerances of 22.5 decimal degrees (i.e., the specified direction $\pm 22.5^\circ$). The maximum of 300 km allows exploration of spatial variation over a large regional scale – it is large enough to assess contrasts between the North and South and between the East and West but small enough to ensure a sufficient number of pairs are included in the calculation of variograms for all directions. The directional variograms for most log-ratios suggest that spatial variation is not constant in all directions. The most notable trend was for larger semivariances (especially at large lags) in the 135° direction (i.e. north west to south east), and the 90° direction (east to west) in relation to the 45° direction (north east to south west), with semivariances at larger lags for the 0° direction also being relatively small. This points to more pronounced differences between the North West and the South East (and between the East and West) than between the North East to South West or between the North and South. These patterns are apparent for both 2001 and 2011. Models fitted to most of the directional variograms would have similar ranges, but different total sills; this is termed zonal anisotropy (Webster and Oliver, 2007). This suggests that the scale of spatial variation is similar in all directions but that the magnitude of the variation differs, for some log-ratios, by direction. There is no simple distinction between, for example, age and socio-economic log-ratios, or consistency within the selected socio-economic

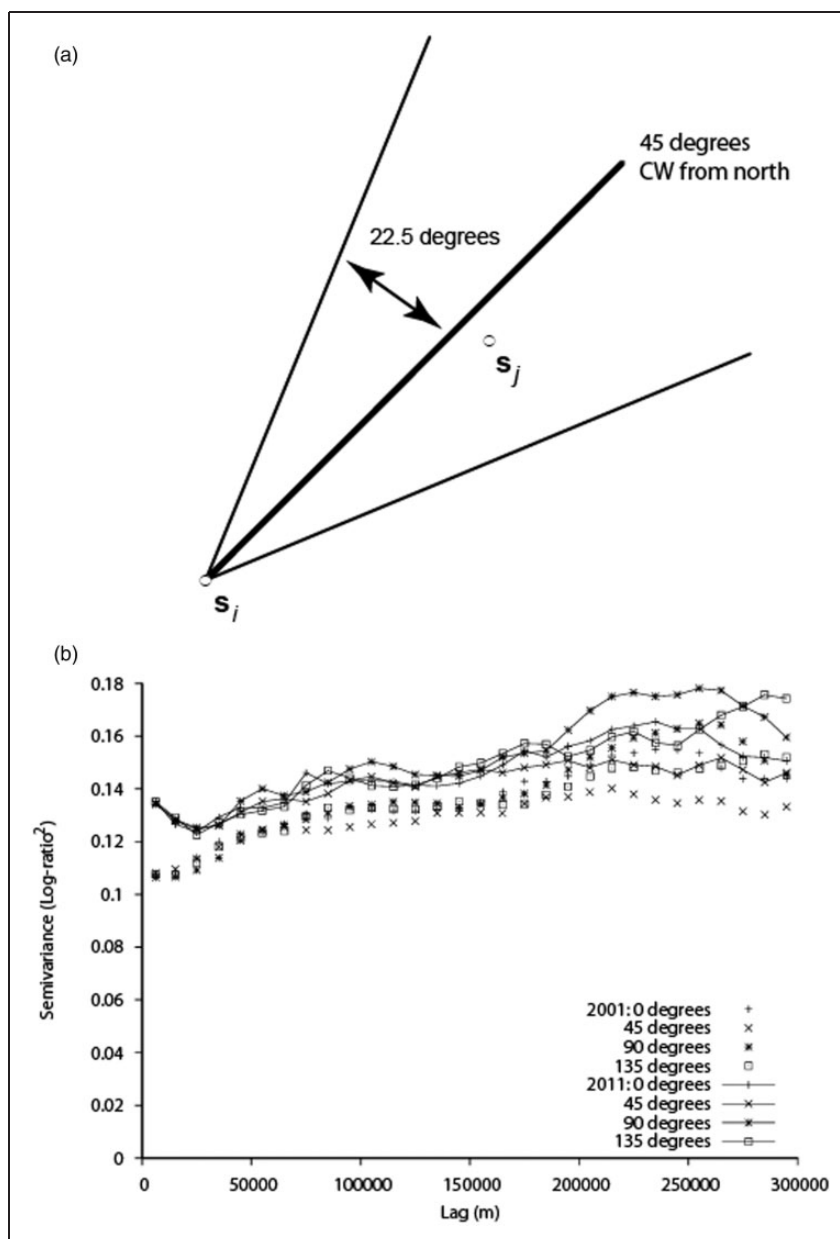


Figure 3. (a) Selection of paired data for directional variogram estimation and (b) directional variogram for LLTI log-ratios for 2001 and 2011.

CW: clockwise.

log-ratios, in terms of the nature of directional variation and thus the directional variograms do highlight distinctive spatial profiles for individual log-ratios. The directional variograms for Tenure1, Tenure2, NS-SEC1 (not shown) and Employ (the latter two variograms have similar forms, as did the equivalent omnidirectional variograms) all indicate a reduction in variation for most (and, in some cases, all) distances and directions. The directional

variograms for NS-SEC2 (not shown) and Qual show some similar features, with ‘dips’ and ‘spikes’ in the semivariances for some lags and directions. However, the directional variograms for NS-SEC2 suggest a reduced variation in some directions, while the equivalent for Qual suggest a small increase in variation for some directions. The variograms for the age log-ratios do not exhibit strong directional effects. It is worth noting that London has a marked impact on the form of variograms in some directions and for some log-ratios given the large number of zones included and the contrast between the North and South and between the North West and South East would likely be enhanced if London were removed from the analyses.

Directional variograms are limited in that semivariances only for selected directions are shown and a more flexible alternative is the variogram map. The variogram map is centred on 0, 0 and the semivariances are binned into the grid cell in which \mathbf{h} is located. That is, the map shows for each cell the average semivariance for the distance and direction represented by the cell (Bivand et al., 2008). Figure 4 shows variogram maps for 2011 and variogram map values for 2011–2001 for three selected log-ratios, Qual, Tenure2 and LLTI. The variogram map for LLTI in 2011 corresponds to the directional variogram in that the largest semivariances are for approximately the 90° (east–west) and 135° (north west–south east) directions and smallest for the 45° direction (north east–south west). For LLTI, for example, differences have increased on average, but they have increased more in some directions than in others – with a large increase in the 135° direction. This suggests that the South and East and the North and West are becoming more dissimilar in terms of LLTI. Fitting a first-order polynomial trend model to the LLTI log-ratios supports visual interpretation of the map of percentages (Figure 1(b)) in suggesting that the LLTI proportions were larger in the North and West than in the South and East in both 2001 and 2011 – so, the trends indicated by the directional variograms and variogram maps reflect these larger LLTI values in the North and West than elsewhere. However, for most log-ratios (despite an increase in differences over most directions and distances for some log-ratios), the differences between, for example, the North West and the South East have remained fairly stable relative to the differences between the North East and the South West. For Tenure2, the variogram maps indicate decreasing differences between all directions. The variogram maps for Qual, as the directional variograms, do not suggest an increase in differences between regions.

Discussion and conclusions

The analysis demonstrates quantitatively that there are distinct spatial patterns for the selected log-ratios. In common with Voas and Williamson (2000), the population by age is found to be fairly geographically even while the distinct spatial structure of the population by ethnicity is apparent; this is demonstrated by examination of the form of the variograms. These patterns conform to expectation, given the observations made under ‘Directional variograms and variogram maps’, but we now have evidence for the scale of variation and the magnitude of changes between 2001 and 2011. This study builds on previous research by characterising the spatial scales over which the log-ratios are spatially dependent. Log-ratios with clear urban–rural patterns (CarsVans and Ethnicity, for example) have variograms which reflect these structures while for other log-ratios (in particular, the age log-ratios, Qual and LLTI), there are less obvious structures over the scales considered. Directional variograms and variogram maps offer a richer perspective on the spatial structure of the log-ratios. For all log-ratios, there is evidence of directional variation but it is particularly pronounced for Qual, NS-SEC2, LLTI and Ethnicity.

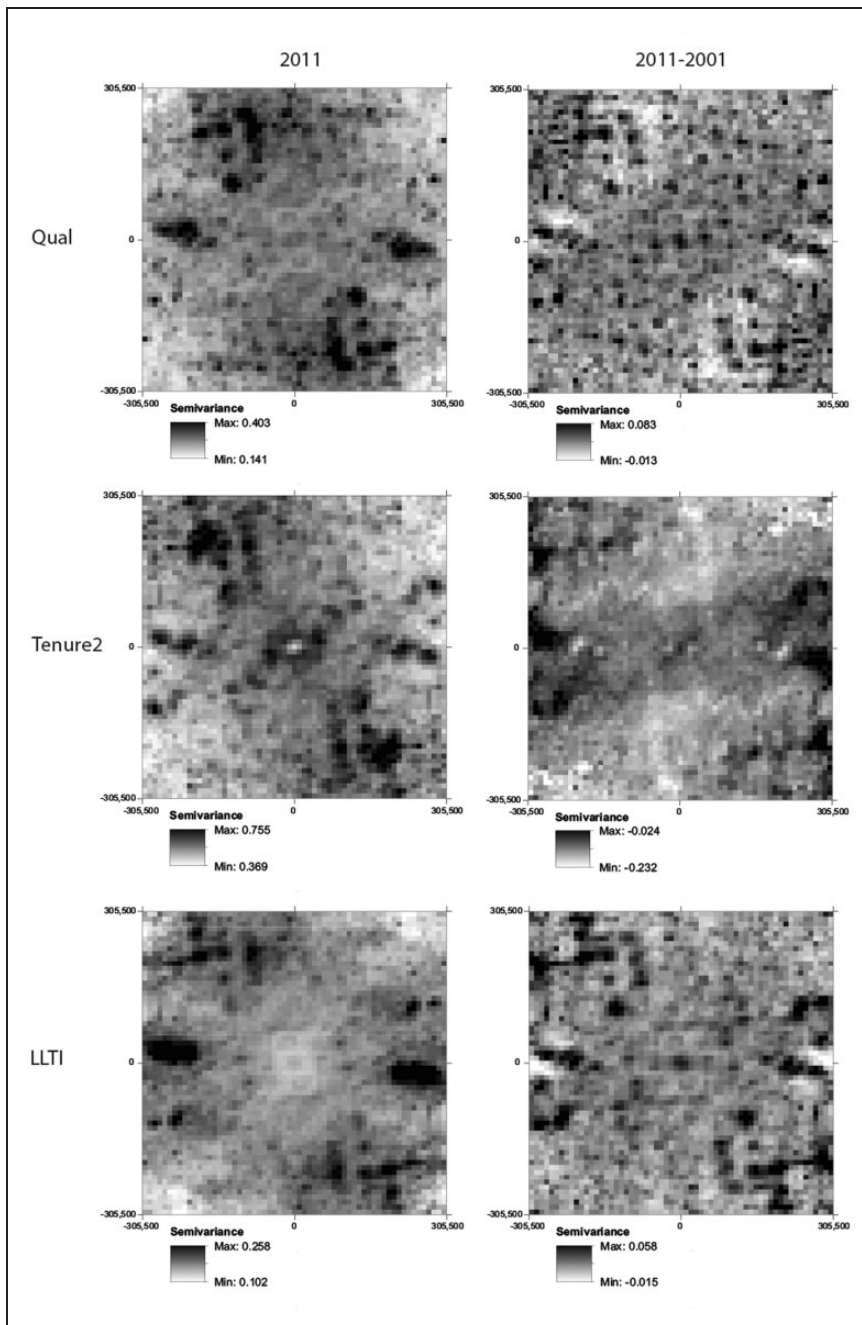


Figure 4. Variogram maps for selected log-ratios for OAs (10 km lag): 2011 and 2011-2001.

Differences between areas have, on average, reduced from 2001 to 2011 with respect to most of the selected log-ratios. But, there are directional variations both in terms of the geographies of the log-ratios in each of the two Census years and also in terms of the magnitude of change. This study has presented the first analysis of differences between

small areas by direction; no previous studies have explored directional variations in a set of population variables without a prior definition of regions such as the North and South. The larger semivariances for the LLTI variograms, and in the variogram maps, in the north west to south east direction and in the east to west direction than in the north east to south west direction support the visual assessment of Figure 1(b) in that generally smaller percentages are found in the South East than in the North and West. There is a suggestion of an increase in differences between the North and West and the South and East in this case, with a proportionally larger increase in semivariances between 2001 and 2011 at larger lags in the north west to south east direction than in the perpendicular direction. Thus, in terms of ill-health, the north–south (or, more correctly, north west–south east) divide in England and Wales appears to be growing. The directional variograms for Qual support the comments of Dorling and Thomas (2004) on a north–south divide in educational attainment. However, as noted previously, the north–south (or, for some log-ratios, north west–south east) differences between 2001 and 2011 are proportionately similar to differences in the perpendicular direction. In this respect, geographical inequalities in qualifications, for example, have increased in all directions (albeit by a relatively small amount) but differences between the North and South (or North West and South East) have not increased relative to differences in other directions.

This analysis is the first to explore directional differences in geographical inequalities using small area data and without prior definitions of regions such as the North and the South. This approach has made it possible, for the first time, to determine how population sub-groups differ between regions and at what spatial scales. The findings highlight increased growth in differences between the North West and the South East in terms of self-reported ill-health, with relatively stable north–south (or north west–south east) differences in other cases. More generally, the findings highlight the limitations of studies of geographical inequalities which do not account for directional variations and studies which define in advance the regions between which differences are to be explored.

There are several obvious ways in which the analysis could be expanded. Increasing the array of variables (as well as, for example, less coarse aggregations by age or ethnicity) and the time frame of the analysis would be particularly beneficial. Including variables which would allow an assessment of geographical differences between the most wealthy and ‘the rest’ (e.g. houses with multiple cars or small households in large houses) would offer another dimension. The present analysis is based on univariate Census outputs and it could be extended to make use of variables based on administrative sources (Norman and Bamba, 2007, discuss the use of sickness benefit data as an indicator of health over small areas). Also, local variograms (see Lloyd, 2012 for an example) could be used to further assess spatial variation in population sub-groups within regions of England and Wales. In addition, expanding the analysis to include the whole of the UK would enable a fuller assessment of the changing relationships between London and the South East and the rest of the country. Finally, making links between the present findings and modelling perspectives which seek to identify the determinants of geographical inequalities would be particularly valuable as a means of better understanding how different the regions of the UK are, and how far these inter-regional differences are changing.

Acknowledgements

The Office for National Statistics are thanked for provision of the data on which the analyses were based. Office for National Statistics, 2001 and 2011 Census: Digitised Boundary Data (England and Wales) [computer file]. ESRC/JISC Census Programme, Census Geography Data Unit

(UKBORDERS), EDINA (University of Edinburgh)/Census Dissemination Unit. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland. The comments of the anonymous reviewers are acknowledged gratefully.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Part of the research on which this paper was based was supported by the Economic and Social Research Council (grant ES/L014769/1).

Notes

1. <http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-wards-and-output-areas-in-england-and-wales/index.html>
2. <http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/output-area-oas-/index.html>

References

- Aitchison J (1986) *The Statistical Analysis of Compositional Data*. London, UK: Chapman and Hall.
- Balabdaoui F, Bocquet-Appel J -P, Lajaunie C, et al. (2001) Space-time evolution of the fertility transition in India, 1961–1991. *International Journal of Population Geography* 7: 129–148.
- Bivand RS, Pebesma EJ and Gómez-Rubio V (2008) *Applied Spatial Data Analysis with R*. New York, NY: Springer.
- Bland JM (2004) North-south divide in social inequalities in Great Britain: Divide in social class inequalities may exist but is small. *British Medical Journal* 329: 52.
- Brown LA and Chung S-Y (2006) Spatial segregation, segregation indices and the geographical perspective. *Population, Space and Place* 12: 125–143.
- Catney G (2013) *Has neighbourhood ethnic segregation decreased? The Dynamics of Diversity: Evidence from the 2011 Census*. Manchester: Centre on Dynamics of Ethnicity.
- Catney G (2015) Exploring a decade of small area ethnic (de-)segregation in England and Wales. *Urban Studies*. In press.
- Catney G and Simpson L (2010) Settlement area migration in England and Wales: Assessing evidence for a social gradient. *Transactions of the Institute of British Geographers* 35: 571–584.
- Doran T, Drever F and Whitehead M (2004) Is there a north-south divide in social class inequalities in health in Great Britain? Cross sectional study using data from the 2001 census. *British Medical Journal* 328: 1043–1045.
- Dorling D and Rees P (2003) A nation still dividing: The British census and social polarisation 1971–2001. *Environment and Planning A* 35: 1287–1313.
- Dorling D and Thomas B (2004) *People and Places: A 2001 Census Atlas of the UK*. Bristol: The Policy Press.
- Egozcue JJ and Pawłowsky-Glahn V (2005) Groups of parts and their balances in compositional data analysis. *Mathematical Geology* 37: 795–828.
- Egozcue JJ and Pawłowsky-Glahn V (2006) Simplicial geometry for compositional data. In: Buccianti A, Mateu-Figueras G and Pawłowsky-Glahn V (eds) *Compositional Data Analysis in the*

- Geosciences: From Theory to Practice*. [Geological Society Special Publications No 264]. London, UK: Geological Society, pp. 145–160.
- Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, et al. (2003) Isometric logratio transformations for compositional data analysis. *Mathematical Geology* 35: 279–300.
- Fahmy E, Gordon D, Dorling D, et al. (2011) Poverty and place in Britain 1968–99. *Environment and Planning A* 43: 594–617.
- Filzmoser P, Hron K and Reimann C (2009) Univariate statistical analysis of environmental (compositional) data: Problems and possibilities. *Science of the Total Environment* 407: 6100–6108.
- Gardiner B, Martin R, Sunley P, et al. (2013) Spatially unbalanced growth in the British economy. *Journal of Economic Geography* 13: 889–928.
- Goovaerts P (2008) Kriging and semivariogram deconvolution in the presence of irregular geographical units. *Mathematical Geosciences* 40: 101–128.
- Goovaerts P, Jacquez GM and Greiling D (2005) Exploring scale-dependent correlations between cancer mortality rates using factorial kriging and population-weighted semivariograms. *Geographical Analysis* 37: 152–182.
- Hacking JM, Muller S and Buchan IE (2011) Trends in mortality from 1965 to 2008 across the English north-south divide: Comparative observational study. *British Medical Journal* 342: d508.
- Johnston R, Poulsen M and Forrest J (2007) The geography of ethnic residential segregation: A comparative study of five countries. *Annals of the Association of American Geographers* 97: 713–738.
- Lloyd CD (2010) Exploring population spatial concentrations in Northern Ireland by community background and other characteristics: An application of geographically weighted spatial statistics. *International Journal of Geographical Information Science* 24: 1193–1221.
- Lloyd CD (2012) Analysing the spatial scale of population concentrations in Northern Ireland using global and local variograms. *International Journal of Geographical Information Science* 26: 57–73.
- Lloyd CD (2014) *Exploring Spatial Scale in Geography*. Chichester: John Wiley and Sons.
- Lloyd CD (2015) Assessing the spatial structure of population variables in England and Wales. *Transactions of the Institute of British Geographers* 40: 28–43.
- Lloyd CD, Pawlowsky-Glahn V and Egozcue JJ (2012) Compositional data analysis for population studies. *Annals of the Association of American Geographers* 102: 1251–1266.
- Lloyd CD, Shuttleworth IG and Wong DW (2014) *Social-Spatial Segregation: Concepts, Processes and Outcomes*. Bristol: Policy Press.
- Martin D, Nolan A and Tranmer M (2001) The application of zone-design methodology in the 2001 UK census. *Environment and Planning A* 33: 1949–1962.
- Massey DS and Denton NA (1988) The dimensions of residential segregation. *Social Forces* 67: 281–315.
- Musterd S (2005) Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs* 27: 331–348.
- Norman P (2010) Demographic and deprivation change in the UK, 1991–2001. In: Stillwell J, Norman P, Thomas C, et al. (eds) *Spatial and Social Disparities Understanding Population Trends and Processes, Volume 2*. Dordrecht: Springer, pp. 17–35.
- Norman P and Bambra C (2007) Incapacity or unemployment? The utility of an administrative data source as an updatable indicator of population health. *Population, Space and Place* 13: 333–352.
- ONS (Office for National Statistics) (2013) 2011 Census analysis, a century of home ownership and renting in England and Wales. Available at: www.ons.gov.uk/ons/rel/census/2011-census-analysis/a-century-of-home-ownership-and-renting-in-england-and-wales/index.html (accessed 23 November 2015).
- Pawlowsky-Glahn V and Olea RA (2004) *Geostatistical Analysis of Compositional Data*. New York, NY: Oxford University Press.
- Pawlowsky V and Burger H (1992) Spatial structure analysis of regionalized compositions. *Mathematical Geology* 24: 675–691.
- Pebesma EJ and Wesseling CG (1998) Gstat, a program for geostatistical modelling, prediction and simulation. *Computers and Geosciences* 24: 17–31.

- Quillian L and Lagrange H (2013) *Socio-Economic Segregation in Large Cities in France and the United States*. [Working Paper Series WP-13-24]. Evanston, IL: Institute for Policy Research, Northwestern University.
- Reardon SF, Farrell CR, Matthews SA, et al. (2009) Race and space in the 1990s: Changes in the geographic scale of racial residential segregation, 1990–2000. *Social Science Research* 38: 55–70.
- Reardon SF, Matthews SA, O'Sullivan D, et al. (2008) The geographic scale of metropolitan racial segregation. *Demography* 45: 489–514.
- Riva M, Curtis S and Norman P (2011) Residential mobility within England and urban-rural inequalities in mortality. *Social Science and Medicine* 73: 1698–1706.
- Simpson L and Finney N (2009) Spatial patterns of internal migration: evidence for ethnic groups in Britain. *Population, Space and Place* 15: 37–56.
- Smith GR (1999) *Area-Based Initiatives: The Rationale and Options for Area Targeting*. Case paper no. 25. London, UK: Centre for Analysis of Social Exclusion, The London School of Economics.
- Stewart MB (2011) The changing picture of earnings inequality in Britain and the role of regional and sectoral differences. *National Institute Economic Review* 218: R20.
- Thomas B, Dorling D and Smith GD (2010) Inequalities in premature mortality in Britain: Observational study from 1921 to 2007. *British Medical Journal* 341: c3639.
- Townsend P (1979) *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living*. Harmondsworth: Penguin Books.
- Tunstall R (2011) *Social Housing and Social Exclusion 2000–2011. Case paper, no. 153*. London, UK: Centre for Analysis of Social Exclusion, The London School of Economics and Political Science.
- Voas D and Williamson P (2000) The scale of dissimilarity: Concepts, measurement and an application to socio-economic variation across England and Wales. *Transactions of the Institute of British Geographers New Series* 25: 465–481.
- Webster R and Oliver MA (2007) *Geostatistics for Environmental Scientists*, 2nd ed. Chichester: John Wiley and Sons.
- Williamson P (2007) The impact of cell adjustment on the analysis of aggregate census data. *Environment and Planning A* 39: 1058–1078.
- Wong D W S (2004) Comparing traditional and spatial segregation measures: A spatial scale perspective. *Urban Geography* 25: 66–82.